Huigen, M. G. A., K. P. Overmars, and W. T. de Groot. 2006. Multiactor modeling of settling decisions and behavior in the San Mariano watershed, the Philippines: a first application with the MameLuke framework. *Ecology and Society* **11**(2): 33. [online] URL: http://www.ecologyandsociety.org/vol11/iss2/art33/

Research, part of a Special Feature on Empirical based agent-based modeling

Multiactor Modeling of Settling Decisions and Behavior in the San Mariano Watershed, the Philippines: a First Application with the MameLuke Framework

Marco G. A. Huigen [1], *Koen P. Overmars* [1], *and Wouter T. de Groot* [1]

ABSTRACT. Land-use system dynamics and demographic dynamics are tightly coupled. In environmental science and studies of changes in land use and land cover, an unequivocal relationship is sometimes found between both systems, especially in coarse-scale studies. To obtain a better understanding of these intermingling dynamics, we formulated an agent-based model, the MameLuke settlement model, that used a deductive approach to investigate these relationships. The model was constructed based on ethnographic histories of farm households in San Mariano, the Philippines. The model was calibrated visually. Although this calibration approach proved to be very inefficient, the model itself still outperformed a random model. The model formulation process and the model outcomes were quite extensively discussed with stakeholders, and the conceptual modeling approach and framework proved to be clear and useful tools for local-scale studies dealing with interacting human and biophysical subsystems.

Key Words: *actor decision-making; agent-based models; ethnic distribution; ethnographic history; land-use dynamics; Mameluke framework; Philippines; population dynamics*

INTRODUCTION

Land-use dynamics and demographic dynamics are tightly interconnected. In environmental science and in studies of changes in land use and land cover, the interaction between these two systems is examined intensively. In statistical studies, an unequivocal relationship is sometimes found between them, especially in coarse-scale studies. For instance, Rudel (1989), Palo (1994), and Rudel and Roper (1997) found a strong influence of demographic pressures on changes in land use and land cover using cross-sectional data. Nevertheless, many studies report statistically insignificant or ambiguous effects of population density on deforestation or land use (Kummer and Sham 1994, Kaimowitz and Angelsen 1998).

One major cause of the ambiguity in these statistical studies is the intermingling through time of processes from the land-use system and the demographic system. The multitude of feedback loops between both systems makes it very difficult to identify a causality between two variables. For instance, if soil erosion is the effect of the overloading of an extensive land-use system, one will find a positive relation between population density and the level of erosion. However, over time, when the population increases and hence the amount of land per capita diminishes, people may begin to invest in terracing and intensify their land-use system; one will then find a negative relation between population density and the level of erosion (Tiffen et al. 1994, Tiffen 1995). Another important element causing this ambiguity is the spatial scale of measurement (Verburg et al. 2003). For instance, at the village level, a crop such as yellow corn may be profitable only on large plots; thus, villages with a relatively high population density may have less corn. However, if yellow corn is produced for large city markets and large-scale piggeries, then at a national level one will find a positive relation between population size and the level of corn production. A third important aspect causing ambiguity is the functional meaning of a variable in the context of land use. For example, if a new road connects poor farmers to markets for arable crops, the farmers may be helped and rely less on slash-

[1] Leiden University

and-burn practices. In this case, one will find a negative relation between the number of roads and deforestation. At the same time, if a new road cuts through the forest and attracts new migrants, the statistical relationship would be positive, with more roads causing more deforestation. This ambiguity is also identified by Pfaff (1999) when he asks the question, "Does population pressure cause deforestation or is deforestation attracting population or is it both at different times?"

To obtain a better understanding of these intermingling variables, we have to include knowledge of the causal processes in studies of changes in land use and land cover. Basically, there are two ways to do this: (1) by enriching the interpretation and analysis of statistical data with causal insights (e.g., Rudel and Roper 1997, Overmars and Verburg 2005) and (2) by setting the causal processes as a starting point in a deductive modeling approach, which is called "generative" modeling by Epstein and Axtell (1996).

Here, we present an agent-based application, the MameLuke settlement model, constructed within the MameLuke framework (Huigen 2004). The source code for both the framework and the model is contained in Appendix 1. This model uses a deductive approach to investigate the relationships between demographic processes and land-use dynamics. It combines demographically related decisions, such as getting married and having children, and land-use decisions, such as settling, resettling, and farm expansion.

The specific contribution of the MameLuke settlement model to this type of generative modeling study is that the structure and content of these decisions were gathered via anthropological and socioeconomic fieldwork. The collected data and "real-life stories" or narrative were transformed into an agent-based computer model within the MameLuke framework following a defined sequence. It is "ethnography first," so to speak, and our aim was to show concretely how this type of study is done and what the results are. Moreover, we aimed to illustrate the three roles of local stakeholders in the modeling process, i.e., in the formulation of the narrative, verification of the modeled processes or agent rules, and discussion of the results.

The MameLuke settlement model has similarities with the Anasazi settlement models (Johnson et al. 2005). In the Anasazi models, households also decide their settlement location while taking several biophysical and social variables into account. These models contribute to a better understanding of the population dynamics and the spatial implications of the demographic dynamics based on field experiences. We also accounted for the role of ethnicity in affecting the location decision of a settler in terms of attraction by and segregation of ethnic groups. This created similarities with the model of Schelling (1971). The simulation of these decisions in a spatial context allowed for a reconstruction of possible ethnic distribution patterns.

Here, we first outline the research methodology. We describe the research area and the computer architecture used to build the model, and give the sequential steps used to transform ethnographies or narratives into a conceptual agent-based model. We then provide a detailed description of this transformation, and hence of the model formulation. Finally, we present the results of the simulations and end with a discussion and conclusion.

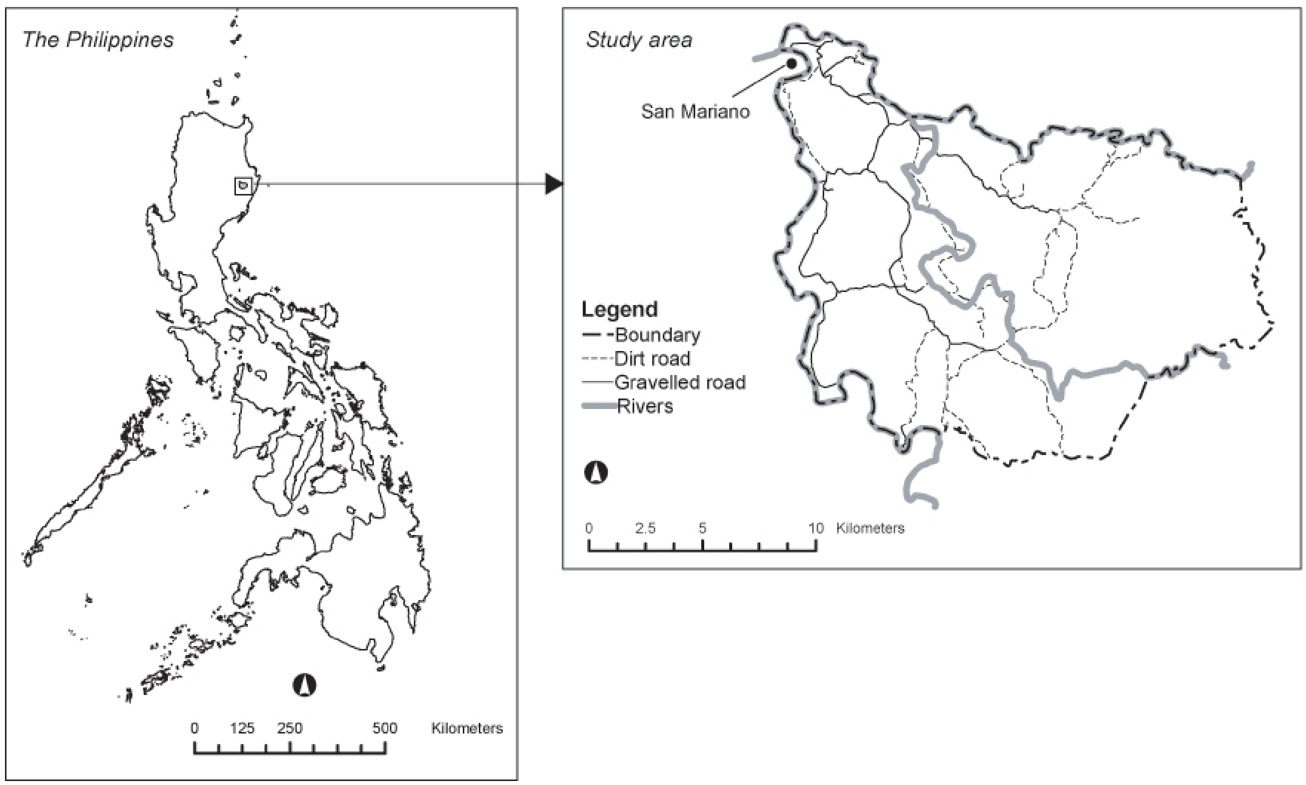
METHODS

Research area

The research area was a watershed in the San Mariano municipality, Isabela, the Philippines (Fig. 1). It reached from the flood plains of the Cagayan River across the Sierra Madre mountains to the Pacific coast. San Mariano municipality contained 36 villages known as "barangays," of which three were classified as urban. Our research area, the watershed of the Disabungan River, which was enclosed by the Disulap River to the north and in part by the Pinacanauan River to the west, consisted of 13 barangays (Fig. 2). Each barangay consisted of 6–11 "sitios." A sitio is an original settlement location. For a detailed description of the physical properties of the research area, see Verburg et al. (2004) and Overmars and Verburg (2005).

A portion of the eastern side of the research area was part of the Northern Sierra Madre Natural Park (NSMNP). The park was established in 1997 to protect and conserve the large number of endemic flora and fauna, such as the Philippine Eagle, *Pithecophaga jefferyi*; the Isabela Oriole, *Oriolus isabellae* (van Weerd and Hutchinson 2004); and the Philippine crocodile, *Crocodylus mindorensis*

Fig. 1. San Mariano research area (P. H. Verburg, *unpublished manuscript*).



(van Weerd and van der Ploeg 2004). An important aspect of the park was that several indigenous communities, such as the Agta (Persoon et al. 2004) and Kalingha, inhabited the area.

Until the late 1980s, a large part of the research area was under logging concessions or timber license agreements. The logging industry started deforesting the area in the early 1950s, and this continued until the late 1980s. The economic boom attracted thousands of people from all over the Philippines. Today, the San Mariano municipality is a melting pot of approximately 30 ethnic identities with their own agricultural practices. The most dominant ethnic identity is the Ilocano group (> 50%); the second largest is the Ybanag (Table 1).

MameLuke framework

The MameLuke framework extends the RePast software framework for agent-based simulation (Minar et al. 1996, Collier 2001). The MameLuke framework is designed at a level of abstraction that allows for the computer implementation of a wide range of models, each of which contains its own specification of the framework, depending on the desires of the modeler. One model may aim to simulate and understand long-term migration flows toward an agricultural frontier, for instance, whereas another may be focused on middle-term land-use responses to extreme events, such as a typhoon. The MameLuke framework simulates discrete time steps on a two-dimensional grid. The time scale is flexible; the user may construct agent rules for various time scales. As in all true agent-based simulations, social agents, who are called actors in the framework, interact with each other and with their environment.

Fig. 2. Barangays in the San Mariano research area. Daragutan East, Disulap, and Ueg were excluded from the analysis because they extended outside the research area. Color gradations indicate different barangays.

Table 1. Actual and simulated ethnic proportions in the San Mariano research area (empirical data from 1997 vs. simulated outcomes of batch 2 [batch 20031018_2; n = 103 runs]).

			Proportion of ethnicity in research area (%)	
Ethnicity	Origin	Main reason for migration	Empirical data (1997)	Average simulation results (1999)
Ilocano	Migrated from Ilocos province	New land	53	55
Ybanag	Cagayan River watershed	Land and labor	24	27
Kalinga	Inhabited San Mariano before 1900	Local	4	2
Ifugao	Ifugao province, Quirino province	Land	3	8
Tagalog	Southern provinces of Luzon	Labor	3	6
Agta	Inhabited San Mariano before 1900	Local	1	2
Other			9	0

The MameLuke framework offers the modeler a generic format to implement this interacting behavior of the modeled entities. This generic format is a hybrid of the traditional computer scientific belief–desire–intention (BDI) architecture (Rao and Georgeff 1995), the agent–group–role (AGR) architecture (Ferber and Gutknecht 1998, Rouchier et al. 1998, Kendall 1999), and behavioral decision and action models (Fishbein and Ajzen 1975, de Groot 1992).

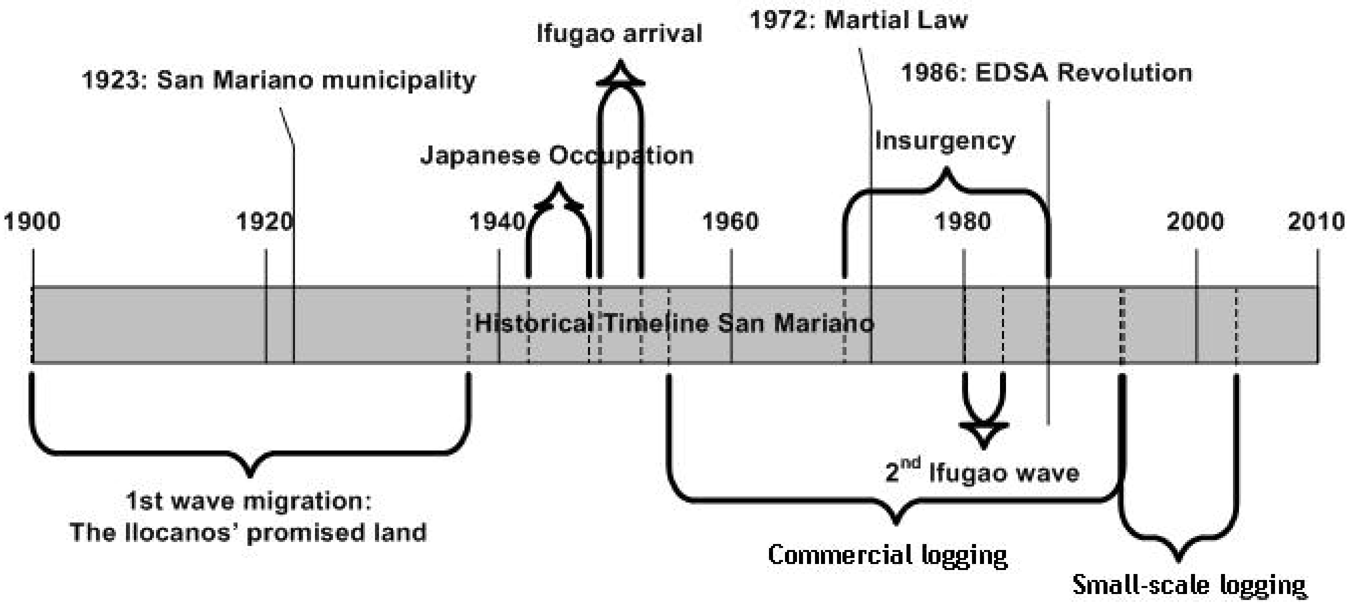
In the framework, the behavioral model of a social agent, i.e., the agent's rules, is structured in potential option paths (POPs) and potential option nodes (PONs). Formally, a PON is a transaction interface between an initializing agent and a recipient agent. A POP defines a sequence of PONs. As a group, the POPs represent a theoretical construct of agent behavior and decisions that the framework user wishes to explore.

A potential action in the mind of a farmer is often a complex of spatial and social interactions. For example, the potential action to grow corn (DoCorn) consists of these subactions: (1) buy seed at store, which is an interaction with the social actor "store"; (2) sow seed, cultivate crop, and harvest yield, which are all separate interactions with a location owned by the farmer; and finally, (3) sell harvest on market, which is a social interaction with the actor "market." To cope with such behavioural complexes in the framework, we differentiated between a POP and a PON. In the above-mentioned example, DoCorn is defined as the POP, which is composed of the PONs BuySeed, SowSeed, etc.

During a time step in a simulation, every agent in the framework may execute its POPs. An agent can only execute those POPs that are available to it, which is determined by the identity groups to which the agent belongs. Hence, every POP contains a list of agent categories that may execute this POP; if an agent is a member of the identity group, all connected POPs are available.

During a simulation, the framework holds all the POPs that can be used. At each time step of the simulation, an agent will formulate its possible actions. These possible actions are a result set, i.e., a list of POPs, of three sequential steps. First, the agent builds a list with implementable options, i.e., available POPs, based upon its categories. Second, the agent puts these available POPs in a specific order based on his motivational structure, which

Fig. 3. Historical timeline for San Mariano.



includes needs, desires, and experience. These ranking algorithms, which represent the internal intelligence of the actor, are beyond the scope of this paper and are explained in forthcoming work. Third, an agent also needs to meet the POP requirements that have been defined by the user. For example, the minimum age to execute the POP GetMarried is 15 yr. If the agent does not meet the POP requirement, it cannot execute the POP, and the POP is not included in the CurrentPOPs list.

Each POP consists of at least one PON and contains a user-defined PON sequence. If an actor wants to execute a POP, it has to pay initiating costs. When an agent executes a POP, it goes through the predefined sequence of PONs. Every PON may again require initiator costs and recipient costs. If a PON is halted because of a logical irregularity, such as the fact that the initiator is unable to find a recipient or the selected recipient does not like the exchange, the whole POP is halted. The initiating agent still pays the initiating costs and will evaluate the POP negatively.

Thus, the PONs are the building blocks of the POPs, and they contain the rationale or calculus of the agent regarding a certain action or exchange. The POPs allow temporally explicit modeling because they have an activation year and an end-activation year. The framework user can use these to set the period that a POP is active. If the simulation time step is not within these two years, the POP cannot be executed. For example, only an agent that is a member of the category Men during the time period 1900–1950 is allowed to execute the POP GetMarried 1900 1950. Agents that are not of the category Men cannot execute this POP during the period 1900–1950; if this were the only rule, no one would marry after 1950. See Huigen (2004) for a more extensive explanation of these components.

Transformation sequence: from narratives to an agent-based model

The MameLuke framework enables us to transform narratives or life histories from the field into a computerized model. Thus, it bridges a gap between more qualitative sciences such as anthropology and more quantitative sciences such as economics. The transformation of narratives into a computerized model is done using a seven-step approach that has similarities with ethnographic decision modeling (Gladwin 1989):

Fig. 4. Statement of Kalingha elder (respondent 1, field trip 5).

"In that time [1920ties] we were making the kaingin [interviewer: swidden] on the flat area, because there were only little people. In 1925 the Kalingha lived already in here in Libertad area. They [interviewer: We] were working on kaingin, but mostly engaged in fishing, hunting and logging. The Ilocanos that came there later were also doing the water logging. That area before was filled with the bamboo. We would take out the bamboo for houses and sell it, then we would make the land ready for upland rice, camote and cassava." (interview: Kalingha elder: respondent 1; ftrip 5).

1. Collection of secondary data. The objective of this first step is to identify the possible variables and processes that require attention.

2. Formulation of hypotheses regarding the selected processes and variables. The aim is to select one or more processes that seem relevant, and to hypothesize the causal mechanisms.

3. Formulation of interview topic lists and questionnaires. The objective is to arrive at a questionnaire that will verify the significance of the processes and provide the necessary information for the hypothesized processes. The questionnaire is formulated while taking into account the structure of the action-in-context (AiC) methodology (de Groot 1992), which includes "progressive contextualization," i.e., reasoning from the actor outward to other actors that influence its decisions.

4. Gathering of primary data.

5. Data analysis and adjustment of hypotheses. Here, we analyze and transform the gathered data into several conceptual models using the AiC methodology.

6. Formulation of a MameLuke model based on the analysis. In this step, we transform the AiC conceptualization into the MameLuke framework.

7. Implementation and analysis of the MameLuke model simulations.

Data collection

During August–November 2002, oral histories were taken from 44 elders, 19 of them female and 25 male, and semistructured interviews were conducted with key actors and groups to gain insight into factors such as migration history, reasons for migration, reasons for settlement, and ethnic differences. For the oral histories, at least three elders from every barangay were interviewed, with the ethnic identities of these respondents corresponding to the dominant ethnic identities of the barangay. In the oral history interviews, we worked along a timeline marking the most important events: (1) arrival of the respondent's parents, (2) respondent's childhood, (3) Japanese occupation and World War II, (4) after the war, (5) start of logging period, (6) period of martial law, and (7) period after martial law and end of logging.

For the key-actor interviews, we chose respondents who have or have had key roles in the community, such as a barangay captain or a delivery woman,

Fig. 5. Statement of Kalingha elder (respondent 1, field trip 5).

"Before the logging [before end 1950ties] we were afraid with the Agta. They would kill you if you were alone. The Agta have a tradition of "Agnayao", if the trees in the forest were flowering, then they attack the people and they would kill someone.
"Sometimes the Agta would kill our tribes people, we do not know why. They have a tradition to do that. For us Kalingas, we knew that we should not go alone into the forest. We had to stay in the 'sitio'. If they killed one of our people we could not do anything, because they were in the forest and we cannot find them. Also, we are very peaceful".
During logging we meet each other because they work as a guide and also we work in forest. When the Aeta already were tame, we could speak to them, they also came here".
(Kalingha elder, respondent 1, ftrip 5)

and people who had distinct jobs, such as a logging company manager. The group interviews were carried out among the regular farm households. The groups consisted of three generations and were mixed in gender.

MAMELUKE SETTLEMENT MODEL

The meta-narrative

The collected life histories were transformed into a MameLuke model. The actor typology was chosen based on a characterization derived from these life histories. The respondents' decisions regarding their demographic and settlement behavior were translated into computerized heuristic rules that were spatially and temporally explicit. Based on the life histories, we divided the 20th century into seven phases identifying the important periods of change (Fig. 3).

At the beginning of 20th century, the San Mariano area was sparsely populated. In those days, the contemporary urban core was a small settlement of people of the Ybanag tribe, who originally came from the Cagayan River plains and worked on the Spanish tobacco farms along the Pinacanauan River or had small farms.

Deeper in the forest along the riverbanks were the settlements of the Kalinghas (Fig. 4). Although some barter relations between the Ybanag and Kalingha occurred in those days, the differences between these groups were great and interactions were rare. The Ybanag were strongly influenced by the Spanish colonization, and thus by Spanish laws, Catholicism, and the educational system and administrative structure. The Kalinghas remained largely outside the range of Spanish influence. They were considered non-Christians because of their animist practices (Keesing 1962). Another non-Christian ethnicity in the San Mariano area is the Agta (Griffin and Estioko-Griffin 1985). From the early 1900s until recently, the Agta were seminomadic. They lived inside the forest, far from the Kalingha and Ybanag settlements, and only occasionally, during hunting trips, would the Kalingha encounter them inside the forest. In general, the Kalingha did not understand much of Agta life and were somewhat afraid of them (Fig. 5).

Fig. 6. Statement of Kalingha elder (respondent 1, field trip 5).

> "The Kalinghas have a very peaceful style. We traded our land for food and livestock with the Ybanags and later on with the Ilocanos who migrated to our areas. In that time it was not our land anyway, and there still was a lot of land to go to, we just left our houses and went along the river into the forest". (interview: Kalingha elder: respondent 1; ftrip 5).

The Ilocanos, who came to the area in the early 20th century in what van den Top (1998) called the "pioneers" migration flow, have a tradition of agriculture. National government policies encouraged land seekers to go to the Sierra Madre area. The settler family could claim a title of ownership on roughly 20–24 ha of uncultivated land. These pioneers came and settled along rivers and on the land of the Kalingha and Agta. They bought the land and the ancestral rights of the Agta and Kalingha for a token payment (van den Top 1998), e.g., livestock and consumer goods such as salt, oil, and alcohol (Fig. 6).

The second phase was the period of Japanese occupation of the Philippines. Nowadays, there are still many stories about this period. During this period there was little immigration.

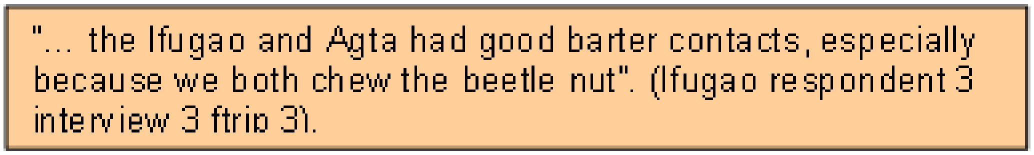
Soon after the war, a group of 40 Ifugao men came to the area, settled in Casala, and brought their families in the years thereafter. The area that they occupied was rather far from the municipality center, and they had to "bunny-hop" over the Kalingha settlements in Libertad and Gangalan. They had to do so because the Kalinghas were already accustomed to the Christian way of thinking (respondent 1, field trip 5) and were no longer leaving or selling their lands. The Ifugao, not intimidated by the stories about the Agta, settled deep inside the forest (respondent 3, field trip 3; Fig. 7). The Ifugao brought with them the knowledge and experience to make rice terraces, and several hillsides in the Sierra Madre are still terraced. This phase also included the start of the corporate logging era. The vast forest area of the Sierra Madre was quickly divided among the influential politicians, and logging companies, often foreign, came in with heavy machinery and started building roads into the forest.

The fourth phase, from the late 1950s to 1979, was the period of the logging boom (Fig. 8). In the late 1950s and early 1960s, San Mariano became accessible via new roads, and in just a couple of years, a thriving timber industry had been established. These economic activities attracted many workers from the southern provinces of Luzon. Although their local dialects differed greatly, they all spoke the Philippine national language, Tagalog, and are hereafter referred to as Tagalog. The abundance of work also attracted many Ybanag laborers. A number of company employees settled permanently in the area.

In 1972, former president Ferdinand Marcos declared martial law, which lasted until 1986. In particular, the first three years were hectic; the San Mariano area was often the center of the battlefield between the government army and the communist rebels of the New People's Army. Twice, for periods that lasted from 1 month to 2 yr, the inhabitants of the area had to be evacuated en masse. After the evacuation, the farm households were allowed to return to their land and houses. In 1980–1985, at the end of martial law, a new group of Ifugao people settled deep inside the forest.

The sixth phase (1986–1992) began in the year that Ferdinand Marcos was replaced, and was the beginning of the end of the logging period. The logging companies were no longer protected by the political regime, and the social and environmental influences in politics demanded a logging ban,

Fig. 7. Statement of Ifugao individual (respondent 3, interview 3, field trip 3).



which was finally approved for San Mariano in 1992.

From that time until the present, the people have been engaged in small-scale logging to supply the local furniture industry, but the economic boom has ended. The San Mariano people quickly transformed into farmers, and it took approximately 3 yr to overcome the economic low.

Transformation of the narratives into an agent-based model

Space and time, actors and input data

The time scale selected for the MameLuke settlement model was a single year because the respondents' memories can be linked to years, either directly or through memorized events, e.g., respondents often said something like, "my parents arrived in San Mariano 3 yr before the Japanese army came." The time extent was 100 yr. The model simulation began in 1900.

The physical world in the model was a representation of spatial aspects that were considered important for the studied system. Their importance was decided based on the stories of the respondents. The physical world was represented by 1-ha grid cells. This size was chosen because our respondents used this unit of measurement. Every grid cell was a location agent with its own characteristics. A location could be a member of one or more agent categories. These agent categories represented GIS raster maps indicating the big rivers, roads, forests, and the slope of the grid cell based on a digital elevation model. See Overmars and Verburg (2005) for details of the preparation of the maps.

The actor categorization in the model was based on the respondents' stories. Each actor category or combination of categories represented a group of people who tended to perform their actions differently from or in the same way as others. In the model, we created individual actors and collective actors. If the collective actor household executed a potential option path (POP), this trickled down to each individual member of that household. The collective actor household was always a member of an ethnic category: Agta, Ilocano, Ybanag, Ifugao, Kalingha, or Tagalog. The individual actors were members of a household. An individual actor inherited the ethnic identity of its household. Furthermore, an individual actor was a member of either the actor category Male or the actor category Female, and individual actors were categorized based on their household position, i.e., father, mother, son, or daughter.

The MameLuke settlement model has four different input data sets: (1) the number of immigrating households per ethnicity per simulated year, (2) the probability of death of an actor given its age, (3) the probability of a married woman becoming pregnant given her age, and (4) the probability of a man getting married given his age. The latter three were used in the POP and potential option node (PON) rule sets of the agents.

The exact number of families of each ethnicity that immigrated or already lived in the area was unknown, but a good estimate could be made by analysis of the contemporary census, historical data, and life stories. Every simulated year, a number of households of a certain ethnicity will arrive within the model world and will try to find a place to settle.

Based on data from the Philippine National Statistics Office (NSO 1995, 2001) and San Mariano census data (Municipality of San Mariano

Fig. 8. Socioeconomic profile of San Mariano in 1981 (Municipality of San Mariano 1995).

> "It was not, however, until San Mariano was made accessible with the construction of the provincial road that links it to the municipality of Naguilian that the former's economic significance was recognized. In the early 1950's, logging magnates began to invest in the timber-rich Sierra Madre mountains. There followed the establishment of sawmills. To date, there are ten sawmills found in the municipality. The extensive logging activities attracted a great number of workers who also brought along with them their respective families." (SEP, 1981)

1997, 2002), the probability of the death of an actor at a given age, the probability of a married woman having a baby at a certain age, and the probability of a man getting married given his age were derived; see Huigen (2004) for more detail. Moreover, the composition of the immigrating households and the ages of their members were determined stochastically. According to the respondents, the families that arrived were relatively young because they were pioneer settlers looking for a new life. In the model, the age of the mother depended on the age of the father, whereas the number of children depended on the age of the mother. The initial age of the father was between 20 and 50 yr. The age of the mother was set at 0–4 yr (stochastic) lower than her husband's. The number of children in an initialized household depended on the age of the mother, such that it equaled a random integer between 0 and a maximum number calculated as the age of the mother minus 16, divided by 2.5. Thus, a mother of 32 yr was stochastically given 0–6 children ([32 – 16]/2.5 = 6.4 children).

The agents' rules: potential option paths and potential option nodes

At the start of each simulation, the model was initiated and the spatial environment was created. At every time step in the model, a given number of households immigrated. Then, the framework randomized the sequential order in which every actor executed its actions.

Table 2 shows the POPs and PONs in the settlement model. The first column gives the name of the POP. The second and third columns specify the period during which a POP can be executed. The fourth column shows the agent categories that were allowed to execute the POP. The fifth column gives the sequence of PONs that was executed within the POP. A POP could be executed by multiple actor categories. A category followed by a plus sign (+) indicates that the agents must be a member of both categories. Thus, for instance, the sixth row indicates that, between the simulated years 1900 and 1931, the Ilocano households would sequentially execute the PONs of the POP SettleAlongRiverNoOtherTribe. The algorithmic contents of these PONs are next explained in detail.

The social actors had several PONs that they executed every time step. These were clustered in the POP called FarmersStep. First, every actor executed the PON GetOlder, which simply increased the age of the actor by 1 yr, and then executed the PON DieOfOldAge. If the actor's probability of death depending on its current age was lower than the value stored in a table of values derived from the NSO (1995), the actor was deleted from the simulation.

Driven by land scarcity in their own province and advised by other pioneer settlers, the early migrants arrived in the San Mariano area on temporary bamboo rafts and moved upstream to the location

Table 2. Potential option paths (POPs) and potential option nodes (PONs) contained within the POPs used to construct the MameLuke settlement model.

POP name	Start year	End y-ear	Agent category	PONs
FarmersStep	1900	9999	Masculine feminine	GetOlder DieOfOldAge
GetMarried	1900	9999	son	GetMarriedToSameEthnicity
GetChild	1900	1988	mother	GiveBirth
GetChild	1989	9999	mother	GiveBirthWithControl
SettleAlongRiverNoOtherTribe	1900	1951	household+kalingha	GotoSameEthnicityGroupMember SettleAlongRiverNoOtherTribes
SettleAlongRiverAmongTribe	1900	1931	household+ilocano	GotoSameEthnicityGroupMember SettleAlongRiverNoOtherTribes
SettleAlongRiverAmongTribe	1900	1931	household+ybanag	GotoSameEthnicityGroupMember SettleAlongRiverNoOtherTribes
SettleAlongRiverAmongTribe	1932	1951	household+ilocano	GotoSameEthnicityGroupMember SettleAlongRiverNotTooManyOtherTribes
SettleAlongRiverAmongTribe	1932	1951	household+ybanag	GotoSameEthnicityGroupMember SettleAlongRiverNotTooManyOtherTribes
SettleAlongRiverAmongTribe	1900	1951	household+ifugao	GotoSameEthnicityGroupMember SettleAlongRiverNoOtherTribes
SettleAlongRiverAmongTribe	1900	1951	household+tagalog	GotoSameEthnicityGroupMember SettleAlongRiverNoOtherTribes
SettleAlongRoadNoOtherTribe	1952	9999	household+agta	GotoSameEthnicityGroupMember SettleAlongRoadNoOtherTribes
SettleAlongRoadNoOtherTribe	1952	1963	household+ilocano	GotoSameEthnicityGroupMember SettleAlongRoadNoOtherTribes
SettleAlongRoadNoOtherTribe	1952	1963	household+ybanag	GotoSameEthnicityGroupMember SettleAlongRoadNoOtherTribes
SettleAlongRoadNoOtherTribe	1952	1979	household+kalingha	GotoSameEthnicityGroupMember SettleAlongRoadNoOtherTribes
SettleAlongRoadAmongTribe	1964	9999	household+ilocano	GotoSameEthnicityGroupMember SettleAlongRoadNotTooManyOtherTribes
SettleAlongRoadAmongTribe	1964	9999	household+ybanag	GotoSameEthnicityGroupMember SettleAlongRoadNotTooManyOtherTribes
SettleAlongRoadAmongTribe	1952	9999	household+tagalog	GotoSameEthnicityGroupMember SettleAlongRoadNotTooManyOtherTribes

(con'd)

SettleAlongRoadAmongTribe	1980	9999	household+kalingha	GotoSameEthnicityGroupMember SettleAlongRoadNotTooManyOtherTribes
SettleAlongRiverNoOtherTribe	1900	1975	household+ifugao	GotoSameEthnicityGroupMember SettleAlongRiverNoOtherTribes
SettleAlongRoadAmongTribe	1976	9999	household+ifugao	GotoSameEthnicityGroupMember SettleAlongRoadNotTooManyOtherTribes
ExpandFarmLand	1900	9999	household	ExpandFarmLand
ReSettle	1935	1935	household+kalingha	ReSettle
CeaseHousehold	1900	9999	household	CeaseHousehold
SaveActorStats	1900	9999	actor	SaveActorStats

of people they knew from their original place. From there, they found their own lots and sometimes registered them at the municipal office. The mode of transport did not change until the roads were built. In the decision for settlement location, the five most important factors indicated by the respondents were: (1) the proximity of co-members of the ethnic group; (2) the availability of a river, e.g., for household purposes and transport back to their place of origin during fiestas; (3) the slope of the land; (4) the soils; and (5) the peace and order situation. Similar findings were reported by Moonen (1998).

Every migrant family, whether Ilocano, Ybanag, or Ifugao, indicated that in the search for a good place to settle, it was important for them to have neighbors they knew directly or indirectly. This ethnic attraction factor was translated in the PON GotoSameEthnicityGroupMember. In the model, the search for a good location to settle began at a settlement location of a stochastically chosen member of the same ethnicity. Once the household actor arrived there, it executed the next PON. If this next PON name contained AlongRiver, the household proceeded to the nearest river location and searched for a location along that river to settle within a Moore-neighborhood. The Moore-neighborhood size depended on the household attribute Actorsview, which differed from ethnicity. If a household had an Actorsview of 3, the search neighborhood was a square with a side of 7 grid cells ([Actorsview × 2] + 1). If it saw a location within that neighborhood that was available and the slope was < 5%, the household would build a farm. If the household did not see a location that fit these requirements, it would move in a random direction to the nearest river location in which it had not yet been and repeat the above process.

The logging boom changed this pattern dramatically. Probably the most important change with regard to the settlement process was road construction. Roads made travel along the river unnecessary. After the logging boom, if the PON name contained AlongRoad, the household executed the same actions as described above, except that the category River was replaced by the category Road.

In addition to the ethnic attraction factor, ethnicity also had a repulsion effect (Fig. 9). The Kalingha moved deeper into the forest, splitting into two groups, one of which was in the research area, once their original location close to the present-day town of San Mariano became too crowded with Ybanags and Ilocanos. To allow the simulation of this repulsion, we added a POP called ReSettle. This POP allowed Agta and Kalingha actors to evaluate and leave their current settlement location. According to the respondents, it took some time for the Ilocanos and Ybanags to become used to each other. The earliest Ilocano settlers kept a good distance between their houses and the Ybanag settlements. The relationship between the Ilocano and Ybanag improved through the years. The newly arrived Ilocano learned to speak Ybanag and started trading their produce. According to the respondents, the physical distance between the Ybanag and

Fig. 9. Statement of Ilocano individual (respondent 12, field trip 3).

> "We [Ilocano] did not speak their [Ybanag] language and we have other traditions. We work much more on the land. The Ybanags would sleep until we would already go to our field. We already did all our morning chores and already have our breakfast." (Ilocano respondent 12, ftrip 3)

Ilocano lands, which ranged from 500 to 1000 m, was much smaller than the distances with the Kalingha people (> 2 km).

To simulate these initial distances between settlements of different ethnic groups, every ethnic group was assigned an attribute called EthnicPush. A household EthnicPush of 4 indicated that ethnic repulsion was active within a square with sides of 9 grid cells ([EthnicPush × 2] + 1). If a PON name contained NoOtherTribes, the actor would continue searching for a location that did not have actors of another ethnic group in its vicinity. If the PON name contained NotTooManyOtherTribes, the household would keep searching for a location that had a majority (> 60%) of its own ethnic background in the vicinity.

If a household had settled, it would try to expand its farmland (POP ExpandFarmLand). A household would expand its farm until the maximum size was reached or land was no longer available. In the model, the maximum size of farmland was an attribute of the ethnic category. The household expanded by executing a spatial search algorithm. The algorithm began at the household's settlement location and used a von Neumann neighborhood to check whether a neighboring location fit the requirements for farm transition. The von Neumann neighborhood was used instead of a Moore neighborhood because, in reality, the farm plots were not often connected diagonally. The algorithm chose the neighboring location with the lowest slope because the respondents indicated that such land had the most value.

In the MameLuke settlement model, a male and a female could get married and form a household. If a male and a female actor married, the farmland that they already possessed by inheritance would become the property of the household. In the POP GetMarried, an actor of the category Son (15 < actor age < 45) would start a search for an actor of the category Daughter that was 0–4 yr younger, and would ask her to marry. A Daughter actor < 15 yr of age would not accept the marriage proposal. The probability of a man getting married at a certain age was derived from the San Mariano 1997 census (Municipality of San Mariano 1997). In the current model, an actor would only search for an actor with the same ethnicity (PON GetMarriedToSameEthnicity).

Women in the MameLuke settlement model had children via the execution of the POP GetChild. If a woman had a baby, the child automatically became a member of the household. A newborn actor had an equal probability of becoming a member of the actor category Male or Female and was respectively added to the category Son or Daughter. One assumption of the model was that only married women would have babies. This assumption was reasonable in the rural Catholic research area. Another assumption in the model was that a woman only had one child per pregnancy. The probability of having a baby was dependent on the age of the woman. The probability that a married woman would have a baby in a particular year was derived from census data collected from approximately 14,000 people in the research area in 1997 and 2002 (Municipality of San Mariano 1997, 2002). It was assumed that this probability was the same for the whole period (1900–2000).

In 1989, a government program started promoting birth control. The female respondents indicated that contemporary young women do not want as many

Fig. 10. Statement of elderly Ilocana individual (respondent 3, field trip 4).

> "Nowadays the youngsters only have three or four children, but in the time I was young we did not have a choice, we would get married at an earlier age and have babies all the time". (Ilocana (1930) respondent 3, ftrip 4)

children as before. According to them, nowadays it is a hardship to take care of four children. The actors in this model did not reason about the number of children. From 1989 onward, the POP GetChild had the PON GiveBirthWithControl, which was similar to the PON GiveBirth, with the restriction that a mother who already had four or more children had only a 10% probability of executing the PON (Fig. 10).

In the POP CeaseHousehold, the execution of the PONs arranged that, when the heads of a household had both died, the possessed land would be equally divided among the children. The gender of the child did not influence the partitioning. This equal division of ownership of the land was practiced by all the ethnic groups in San Mariano. The model randomly divided the land among the children, independent of slope or water availability. If a household was childless, the land became available for occupation by other households.

Finally, the SaveActorStats POP fired a PON that allowed an agent to write its attributes and characteristics to a text file for data analysis purposes.

RESULTS

We performed five simulation experiments. Each experiment contained a number of model runs, which together are called a batch. In batch 1, we conducted 200 runs using only the demographic potential option paths (POPs) to validate the demographic parameter settings. In batch 2, we included the ethnic and spatial aspects of the actor's behaviors; this batch was the complete model. In batch 3, we ran approximately 80 simulations using a different parameter setting than in batch 2; we included more immigrant households and a lower value for the EthnicPush attribute for the Ilocano and Kalingha ethnicities. The main purpose of including batch 3 was to better understand the model dynamics. Batch 4 was a random simulation in which actors built a farm without ethnic influence. It included the same demographic settings as batch 2. The parameter settings of batch 5 were similar to those of the second, and additionally, this batch collected data describing the population increase per barangay per year (Table 3).

We next discuss the calibration and verification of the model. In the calibration of the model, we adjusted the parameters and variables until the model displayed satisfactory preliminary results for the separate model components and as a whole. For instance, we adjusted the total number and ethnic proportions of immigrating households to arrive at a total population size with the correct ethnic proportions corresponding to empirical data. In the verification of the model, we checked the various modules to see if they performed as intended, and whether the distinct components were created and combined logically according to the field-level stakeholders and external experts.

We also provide the static results and validation of these results. In the validation of the model, we measured the model performance. The original experimental design, consisting of batches 1–4, assumed that empirical spatial data for validation purposes were available at the "sitio" level. Unfortunately, the quality of the collected data proved too low for such purposes, and secondary data at a coarser level had to be used. Batch 5 was additionally designed for validation using the available coarse-level data. For this reason, we

Table 3. The experimental setup, indicating whether specific types of potential option paths (POPs) were activated in a batch.

Batch number†	Number of runs	POPs included		
		Demographic	Ethnic	Spatial
20031018_1	200	Yes	No	No
20031018_2	116	Yes	Yes	Yes
20031018_3	~80	Yes	Yes	Yes
20031018_4	~40	Yes	No	Yes
20031018_5	44	Yes	Yes	Yes

†Refers to an internal code in the model.

mainly reported static results from batch 5. Static results refer to model outcomes at a defined simulation time step. Here, we used the results from simulation year 1999 for the analysis. Finally, we provide the dynamic results of the experimental simulations. Dynamic results refer to model results through time; hence, they describe the processes within the model as a result of interactions among the actors.

Model calibration and verification

For the demographic actor PONs in the modeling process, we used both calibration and verification techniques iteratively to arrive at a realistic model implementation. The main demographic input parameters calibrated were the number of immigrant households per year (Fig. 11), the number of children born, and mortality. First, we calibrated the input data to achieve the following outcomes: (1) the simulated total number of people per year was equal to that of the secondary empirical data; (2) the total number of farm households matched the statistical data; (3) the population age distribution was realistic; (4) the distribution of the average number of people per household was equal to that of the census data, e.g., approximately 5.2 heads per household in the 1997 census (Municipality of San Mariano 1997); and (5) the proportions of ethnic identities in simulated 1999 should approximate the 1997 ethnic composition. The calibration of these criteria was successful, as shown in Huigen (2004). The average simulated ethnic proportions of 103 runs in batch 2 were compared with the actual ethnic proportions (Table 1).

Second, we calibrated the parameters of the micro-level actor behaviors to arrive at a visually acceptable macro-scale pattern. Hence, we tried to find a model specification with micro-level parameters that resulted in the emergence of the macro-scale pattern that approximated the modelers' visual perception of reality, with the main focus on obtaining the location of both the Kalingha and the Ifugao ethnicities as accurately as possible. The modelers' spatial target as used for the visual calibration is provided (Fig. 12).

The calibrated parameter settings for the four spatially explicit batches that resulted in a satisfactory visual pattern are shown (Table 4). The parameter "land" is the maximum number of plots per farm per ethnicity to be cultivated. The parameter "push" is the preferred distance, i.e., number of grid cells, of an ethnicity from other ethnic identities. In this calibration step, we also had to adjust the activation years of certain POPs.

Fig. 11. Input data for the number of households that immigrated to the research area by ethnicity per 10-yr period.

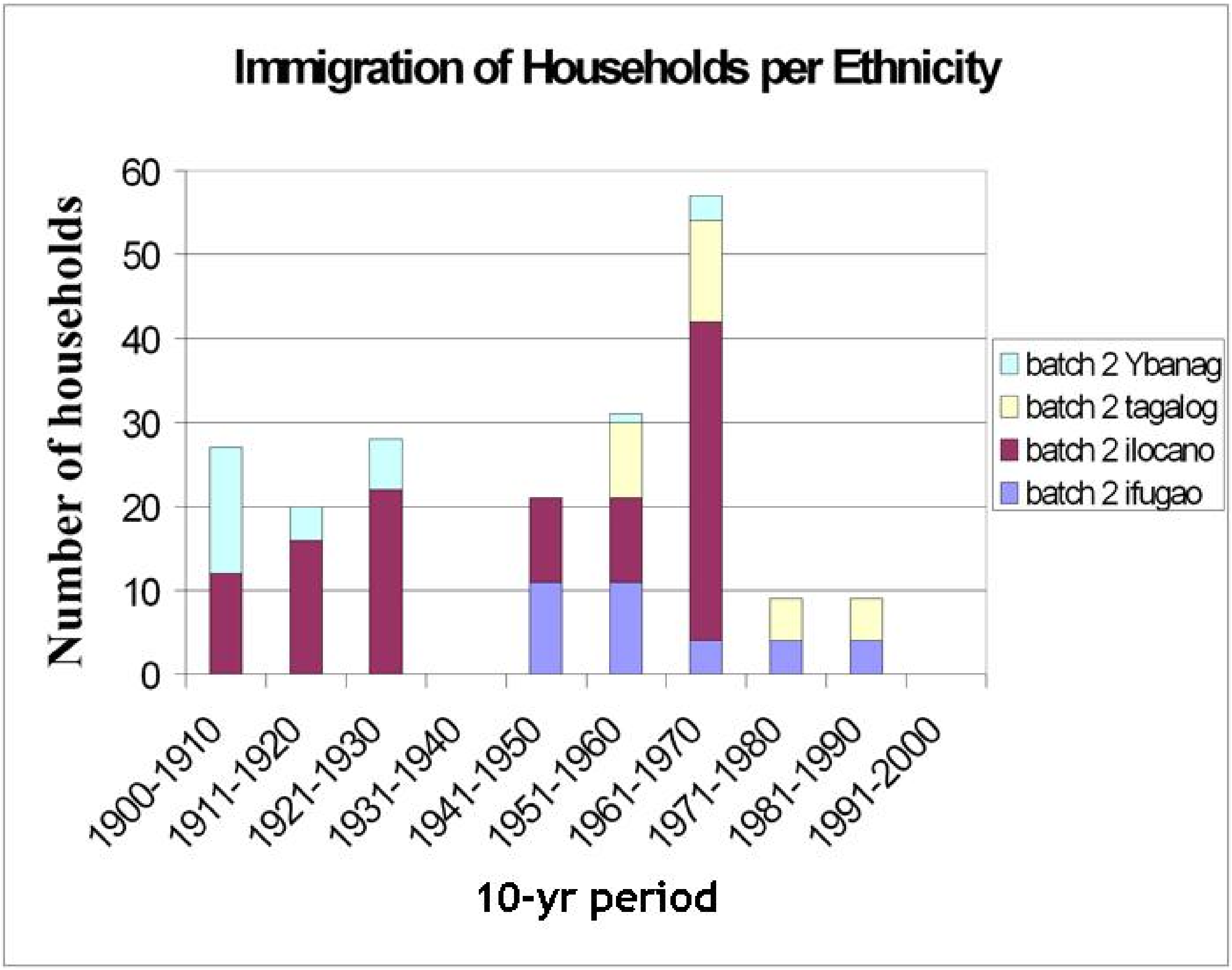


Fig. 12. Locations of Kalingha and Ifugao ethnicities used for the visual model calibration.

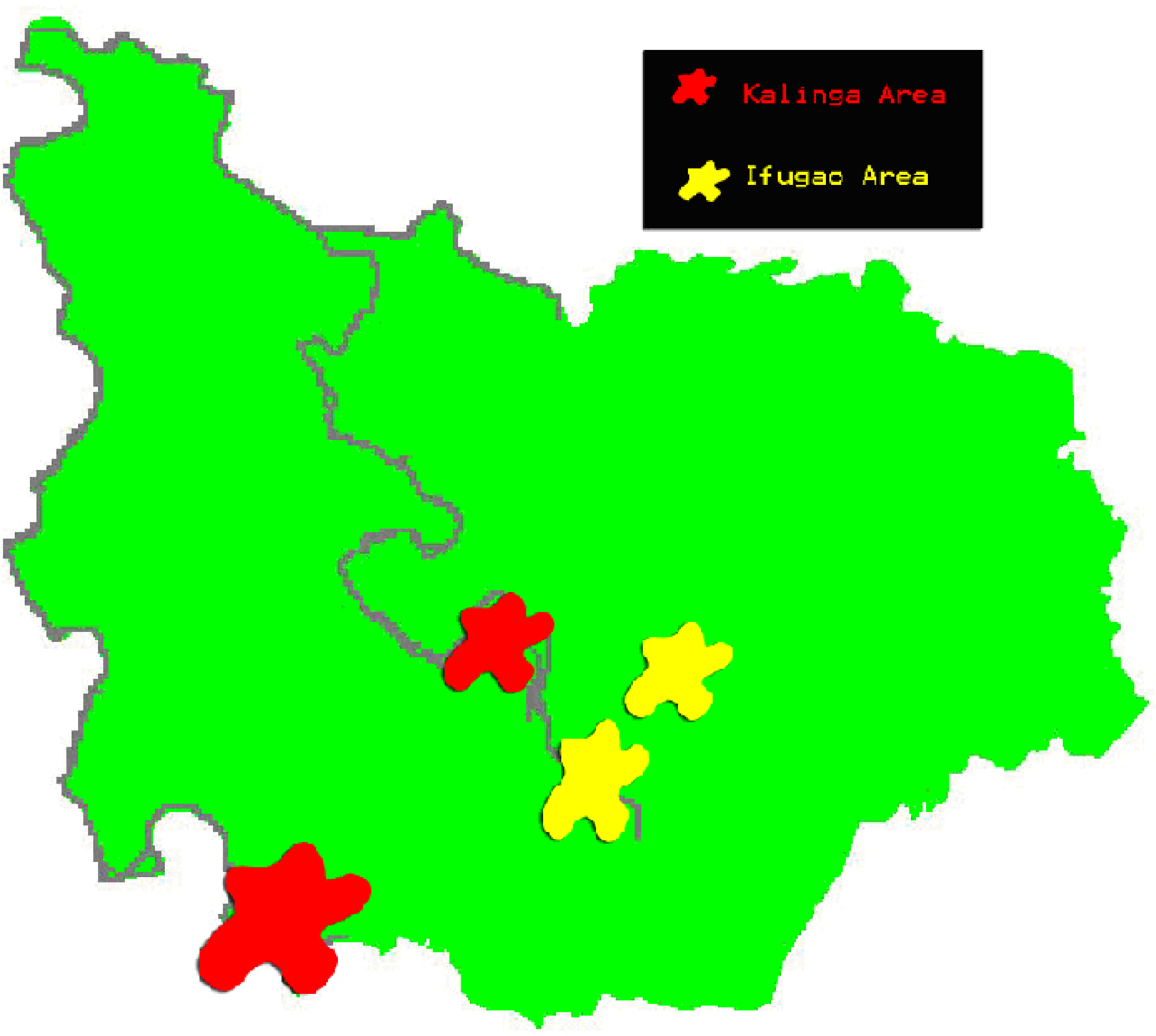


Table 4. Parameters for ethnic push (Push) and the maximum land size per ethnicity (Land) for model simulation batches 2–5.

	Ilocano		Agta		Ybanag		Kalingha		Ifugao		Tagalog	
Batch number	Land	Push	Land	Push	Land	Push	Land	Push	Land	Push	Land	Push
20031018_2	10	10	1	24	8	10	3	20	5	12	3	4
20031018_3	10	8	1	24	8	8	3	14	5	12	3	4
20031018_4	10	N/A†	1	N/A	8	N/A	3	N/A	5	N/A	3	N/A
20031018_5	10	10	1	24	8	10	3	20	5	12	3	4

†Not applicable to this batch.

During the formulation of the conceptual model and the construction of the computer model, we asked scientific experts and local stakeholders to verify the actors' POPs. Several comments were used to adjust previous model versions.

Static results and validation

We visualized one of the model runs from batch 2 to illustrate the working of the model (Appendix 2). Every year, a number of households of a particular ethnicity would establish their farms according to the model POPs.

For every spatially explicit simulation run in batch 5, we generated a map that gives the occupied locations per ethnicity in simulation year 1999. Within a batch, the differences in the spatial ethnic distributions among the runs were visibly large. These differences were caused by the effect of the stochastic elements in the demographic potential option nodes (PONs) because of the exponential behavior of population growth. Each of the 43 runs in batch 5 simulated different model outcomes, varying in total population size, ethnic group proportions, and ethnic group distributions. To obtain the probability that a location was occupied by a certain ethnicity, we combined the total number of model runs and counted the number of model runs in which a location was occupied by an ethnicity per year. This number of model runs was divided by the total number of model simulations to provide the probability that a location was occupied by an ethnicity at a given year for a certain model parameterization. We visualized these probabilities for the Ilocano ethnicity in the years 1910, 1935, 1950, 1960, 1970, 1980, 1990, and 1999 (Fig. 13).

"Best" model runs

Figure 14 shows the two visually "best" model outcomes from batch 5 according to the visual calibrations of the stakeholders and modelers, and illustrates the differences in spatial ethnic-group distribution between the two model outcomes.

These two model outcomes can be considered the best models based on qualitative criteria. To measure the performance of the model quantitatively, the average simulated ethnic proportions of the population per barangay for all runs (n = 43), hereafter referred to as "*sim*," were compared with the empirical ethnic proportions of the population per barangay, hereafter referred to as "*real*," and the random ethnic proportions of the population per barangay, hereafter referred to as "*rand*." The value of *rand* corresponds to an outcome that assumes that the ethnic proportions of the population per barangay are equal to the empirical ethnic proportions of the population in the whole research area. Hence, *rand* is the outcome of

Fig. 13. Dynamic representation of the probability that a location was occupied by an Ilocano household in 1910, 1935, 1950, 1960, 1970, 1980, 1990, and 1999, obtained using the MameLuke settlement model (batch 5).

This figure is animated and may be viewed by clicking here, or by going to: http://www.ecologyandsociety.org/vol11/iss2/art33/figure13.html

the simplest model, assuming a homogenous distribution of the ethnic proportions. Obviously, the interesting question here is whether *sim* performs better than *rand.*

We compared these differences in performance by calculating the validation score; this is the absolute difference between *real* and *sim* or the absolute difference between *real* and *rand.* We calculated these validation scores per model run per barangay per ethnicity for the year 1999.

The first summary of the model performance (Table 5) shows the average validation scores per ethnic group per barangay for *real* (column 4), *sim* (column 5), and *rand* (column 6). Furthermore, it shows the absolute differences between *sim* and *real* (column 7), and between *rand* and *real* (column 8) per barangay per ethnic group. If the *sim* validation score (|*sim* - *real*|) for an ethnicity in a barangay was lower than the *rand* validation score (|*rand* - *real*|), the simulation model performed better than the simplest random model for that ethnicity in that barangay.

The last row in Table 5 displays the sum of the validation scores. Because the sum of the *sim* validation scores was 9% lower than the sum of the *rand* validation scores, a first overall conclusion is that our model performed better than the random model.

These first performance measurements do not account for the effect of population size per barangay. Hence, a barangay with a large population had a greater effect on the performance measurements than did a barangay with a small population. If we applied a population weight correction (Table 5, columns 9 and 10), the sum of the *sim* validation scores was 15% lower than the sum of the *rand* validation scores.

The ethnicity weight (Table 5, columns 11 and 12) corrected for differences in ethnic population size. Without this weight, the errors caused by a large ethnic group, e.g., the Ilocano, influenced the outcome more than those caused by a small ethnic group, such as the Kalingha. The ethnicity weight allowed for an equal influence of each ethnicity. With this ethnicity weight applied, the sum of the *sim* validation scores was 21% lower than the sum of the *rand* validation scores.

The validation scores were the average of all simulation runs in batch 5 (Table 5). However, single runs differed greatly in their validation scores. Figure 15 shows the calculated simulation validity score per simulation run. The red line represents the score of *rand.* The majority of the model runs scored better than *rand* (70%) when only the population weight was applied.

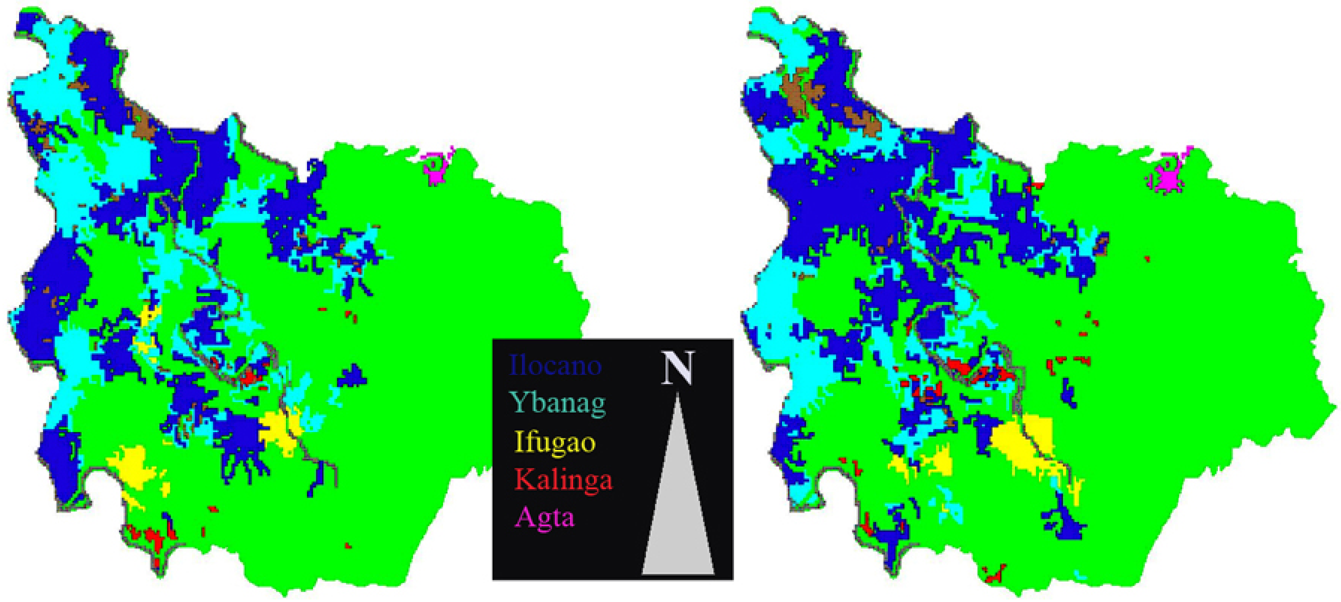
The best simulation run (*sim*_id = 1563327) had a 35% lower total validation score (40.08) than *rand* if the population weight was applied. If the ethnicity weight was also applied, 90% of the runs scored better than *rand.* The best-scoring run (*sim*_id = 3102335) had a 57% lower total validation score than *rand.* Both best-scoring runs are represented spatially in Fig. 16.

Dynamic model results

Dynamic results suggest that, over time, the average plot size per barangay decreased. This occurred most quickly in the barangays closest to San Mariano proper, i.e., nearest to the entry point for immigrating households. Furthermore, the analysis showed that, over time, the average plot size per ethnicity decreased, especially for the Ybanag and Ilocano. The households in Libertad had a low average plot size because many of them were of the Kalingha ethnicity, which was constrained to three locations per household (Table 4).

The main reason for the decrease in average plot size over time was natural population growth.

Fig. 14. Two "best" model outcomes in year 1999 based on the visual calibration criteria (batch 5; left: simulation ID = 9842585, right: simulation ID = 1064722).



Households inherited a certain number of farm locations, which in most cases were smaller than their parents' original farms, yet large enough for them to decide not to migrate to a further settlement. These households had no possibility of expanding their farms because all the land in the vicinity was already occupied. Hence, the same number of locations in, for example, 1960 was occupied by a larger number of households in 1990.

Furthermore, we observed that the randomization of immigrating households per time step had a strong path-dependent effect on the spatial allocation of the ethnic groups. Randomizing the order of immigrating households of different ethnicities created more spread and a diverse macro-scale ethnic composition because the PON SettleAlongRiverNoOtherTribes forced a certain distance between different ethnicities.

If a large number of households arriving in the area happened to have the same ethnicity, the spatial macro-scale pattern was denser, especially at the beginning of a simulation run; the area occupied by that ethnicity had a much larger connected surface. Moreover, a larger actor's view resulted in larger clusters of ethnicities because the households could settle further from the river, and it took longer for the full stretch of a river to be used by the ethnicities. Therefore, a smaller actor's view resulted in a more dispersed allocation of ethnicities along the river.

Furthermore, the higher the ethnic repulsion factor (EthnicPush), the larger the distance between the initial ethnicity settlements and the faster the full stretch of a river was filled. At the beginning of a simulation, it resulted in a high number of small ethnic clusters with a wide spread, and hence a large degree of ethnic segregation. However, in a later phase of the model, it resulted, from a spatial point of view, in a larger degree of ethnic integration. In later periods, the households of different ethnic backgrounds filled the empty spaces among the original clusters.

Finally, the ratio of ethnic repulsion factors (EthnicPush) per ethnicity also had an influence on the ethnicity allocation. If an ethnicity had an equal or smaller factor than the other ethnicities, then, in the case of the Kalinghas, the households were either forced to resettle or, in other cases, stop growing. Table 6 shows the major descriptive findings of the modeling exercise.

Table 5. Empirical (*re*), random (*ra*), and simulated (*s*) average population proportions per ethnicity per barangay from batch 5 (*n* = 43). % stands for percentage of population.

												Summed per barangay					
			Without population weight					With population weight				Without population weight		With population weight			
								Without ethnicity weight		With ethnicity weight				Without ethnicity weight		With ethnicity weight	
Barangay	%	Tribe	*re*	*s*	*ra*	*s-re*	*ra-re*	*s-re*	*ra-re*	*s-re*	*ra-re*	*s*	*ra*	*s-re*	*ra-re*	*s-re*	*ra-re*
Alibadabad	10.7	Ifugao	0.00	0.01	3.53	0.01	3.53	0.00	0.38	0.00	0.11	14.88	60.99	1.59	6.53	0.08	0.90
		Ilocano	67.29	74.71	63.30	7.42	3.99	0.79	0.43	0.01	0.01						
		Kalingha	2.04	0.04	4.16	2.00	26.96	0.21	2.89	0.05	0.69						
		Ybanag	30.67	25.23	29.00	5.44	26.51	0.58	2.84	0.02	0.10						
Binatug	10.4	Ifugao	0.00	2.70	3.53	2.70	3.53	0.28	0.37	0.08	0.10	12.41	17.25	1.29	1.79	0.19	0.15
		Ilocano	58.01	61.72	63.30	3.71	5.29	0.39	0.55	0.01	0.01						
		Kalingha	4.58	0.59	4.16	3.99	0.42	0.41	0.04	0.10	0.01						
		Ybanag	37.01	34.99	29.00	2.02	8.01	0.21	0.83	0.01	0.03						
Casala	7.5	Ifugao	19.95	48.14	3.53	28.19	16.42	2.12	1.24	0.60	0.35	81.93	44.94	6.17	3.38	0.78	0.48
		Ilocano	69.29	28.39	63.30	40.90	5.99	3.08	0.45	0.05	0.01						
		Kalingha	0.00	6.08	4.16	6.08	4.16	0.46	0.31	0.11	0.08						
		Ybanag	10.63	17.39	29.00	6.76	18.37	0.51	1.38	0.02	0.05						

(con'd)

Cataguing	9.5	Ifugao	0.00	0.00	3.53	0.00	3.53	0.00	0.34	0.00	0.10	83.55	136.49	7.96	13.00	0.20	0.50
		Ilocano	2.26	44.53	63.30	42.27	61.04	4.02	5.81	0.06	0.09						
		Kalingha	0.00	0.00	4.16	0.00	4.16	0.00	0.40	0.00	0.10						
		Ybanag	96.76	55.47	29.00	41.29	67.76	3.93	6.45	0.14	0.22						
Del Pilar	5.7	Ifugao	35.08	20.34	3.53	14.74	31.55	0.84	1.80	0.24	0.51	52.63	63.10	3.00	3.60	0.37	0.62
		Ilocano	61.59	50.02	63.30	11.57	1.71	0.66	0.10	0.01	0.00						
		Kalingha	0.00	5.93	4.16	5.93	4.16	0.34	0.24	0.08	0.06						
		Ybanag	3.33	23.71	29.00	20.38	25.67	1.16	1.46	0.04	0.05						
Dipusu	4.1	Ifugao	0.00	0.00	3.53	0.00	3.53	0.00	0.15	0.00	0.04	23.95	9.46	0.99	0.39	0.11	0.09
		Ilocano	62.55	59.47	63.30	3.08	0.75	0.13	0.03	0.00	0.00						
		Kalingha	8.79	0.00	4.16	8.79	4.63	0.36	0.19	0.09	0.05						
		Ybanag	28.45	40.53	29.00	12.08	0.55	0.50	0.02	0.02	0.00						
Disulap	13.3	Ifugao	0.00	0.15	3.53	0.15	3.53	0.02	0.47	0.01	0.13	42.72	51.86	5.66	6.87	0.15	0.40
		Ilocano	89.04	67.87	63.30	21.17	25.74	2.81	3.41	0.04	0.05						
		Kalingha	0.00	0.05	4.16	0.05	4.16	0.01	0.55	0.00	0.13						
		Ybanag	10.57	31.92	29.00	21.35	18.43	2.83	2.44	0.10	0.08						
Gangalan	5.5	Ifugao	0.00	25.63	3.53	25.63	3.53	1.40	0.19	0.40	0.05	3.71	19.04	2.93	1.04	0.47	0.12
		Ilocano	68.88	45.93	63.30	22.95	5.58	1.25	0.30	0.02	0.00						
		Kalingha	8.10	4.21	4.16	3.89	3.94	0.21	0.21	0.05	0.05						
		Ybanag	23.01	24.24	29.00	1.23	5.99	0.07	0.33	0.00	0.01						

(con'd)

Libe-rtad	6.4	Ifug-ao	46.19	19.73	3.53	26.46	42.66	1.70	2.74	0.48	0.78	74.71	139.-21	4.80	8.94	0.71	1.29
		Iloc-ano	18.89	32.95	63.30	14.06	44.41	0.90	2.85	0.01	0.05						
		Kali-ngha	30.95	20.21	4.16	10.74	26.79	0.69	1.72	0.17	0.41						
		Yba-nag	3.65	27.11	29.00	23.46	25.35	1.51	1.63	0.05	0.06						
Mar-annao	7.1	Ifug-ao	0.00	0.00	3.53	0.00	3.53	0.00	0.25	0.00	0.07	33.92	36.79	2.42	2.63	0.06	0.19
		Iloc-ano	81.70	64.74	63.30	16.96	18.40	1.21	1.31	0.02	0.02						
		Kali-ngha	0.00	0.00	4.16	0.00	4.16	0.00	0.30	0.00	0.07						
		Yba-nag	18.30	35.26	29.00	16.96	10.70	1.21	0.76	0.04	0.03						
Pan-ninan	4.6	Ifug-ao	0.00	44.98	3.53	44.98	3.53	2.05	0.16	0.58	0.05	112.-15	61.77	5.12	2.82	0.91	0.41
		Iloc-ano	57.05	28.56	63.30	28.49	6.25	1.30	0.29	0.02	0.00						
		Kali-ngha	33.89	7.47	4.16	26.42	29.73	1.21	1.36	0.29	0.33						
		Yba-nag	6.74	19.00	29.00	12.26	22.26	0.56	1.02	0.02	0.04						
San Jose	11.2	Ifug-ao	0.00	5.63	3.53	5.63	3.53	0.63	0.39	0.18	0.11	75.44	72.51	8.43	8.11	0.40	0.40
		Iloc-ano	99.36	61.84	63.30	37.52	36.06	4.19	4.03	0.07	0.06						
		Kali-ngha	0.00	1.47	4.16	1.47	4.16	0.16	0.47	0.04	0.11						
		Yba-nag	0.24	31.05	29.00	30.81	28.76	3.45	3.22	0.12	0.11						

(con'd)

San Pedro	4.0	Ifugao	0.00	0.00	3.53	0.00	3.53	0.00	0.14	0.00	0.04	49.87	63.80	1.99	2.55	0.05	0.13
		Ilocano	94.98	70.27	63.30	24.71	31.68	0.99	1.26	0.02	0.02						
		Kalingha	0.00	0.00	4.16	0.00	4.16	0.00	0.17	0.00	0.04						
		Ybanag	4.57	29.73	29.00	25.16	24.43	1.00	0.98	0.03	0.03						
Total:						711.87	777.24	52.35	61.64	4.48	5.70	711.87	777.24	52.35	61.64	4.48	5.70

DISCUSSION AND CONCLUSION

The model

The model demonstrated that, from relatively simple micro-scale, i.e., individual and household, rules, a macro-scale structure of ethnic allocation patterns emerged that performed much better than a random model with regard to the validation score; 90% of the model runs performed better than the random model. The model runs that were selected in the calibration of the model as having the best visual pattern scored only marginally better than random settlement (*rand*). The total validation score (59.75) of the first simulation run (*sim*_id = 1064722) showed that it performed 3% better than *rand* with only the population weight applied. The first simulation scored much better than the random model when the ethnicity weight was applied, with a total validation score 35% lower than that of *rand*. The second simulation run (*sim*_id = 9842585), with a total validation score of 63.96, performed worse than *rand* when only the population weight was applied. Even though it performed better than the random model when the ethnicity weight was applied, it was one of the lower ranked simulations. These results suggest that the visual calibration approach does work because both simulations had a higher ranking when the ethnicity weight was applied. However, they also show that the visual calibration was not executed very well. The validation scores per ethnicity per run with both correction weights applied underscore this result; the validation scores were increased mainly by erroneously located Ifugao and Kalingha households.

The results also show that visual calibration is not a preferable technique because both "visually best" runs had low scores with the applied validation approach at the barangay level. The visual calibration focused too much on the location of the Kalingha and Ifugao and disregarded the other larger ethnicities. Therefore, future work on a new settlement model should use a more advanced calibration technique to include the settlement locations of the Ybanag and Ilocano.

When the validation scores were totaled per barangay, *sim* performed better than *rand* for most of the barangays, especially Alibadabad, Cataguing, and Libertad; however, for the barangays Casala, Dipusu, Gangalan, and Panninan, *rand* significantly outperformed *sim*. Therefore, further work on the parameterization and calibration of a new settlement model should emphasize the settlement processes in these areas for better model performance. Application of the ethnicity weight correction made *sim* perform better than *rand* for both the overall model and the individual runs. Likewise, at the barangay level, the performance increased for most of the barangays; however, the validation scores for the barangays Binatug, Gangalan, and Panninan decreased dramatically. These were three large barangays in the middle of the research area. In the current model, the immigrant settlers arrived in these barangays during the simulation years 1915–1925. This finding indicates that a new settlement

Fig. 15. Calculated simulation validation scores with population weight (upper) and with population weight and ethnicity weight (lower; batch 5, *n* = 43).

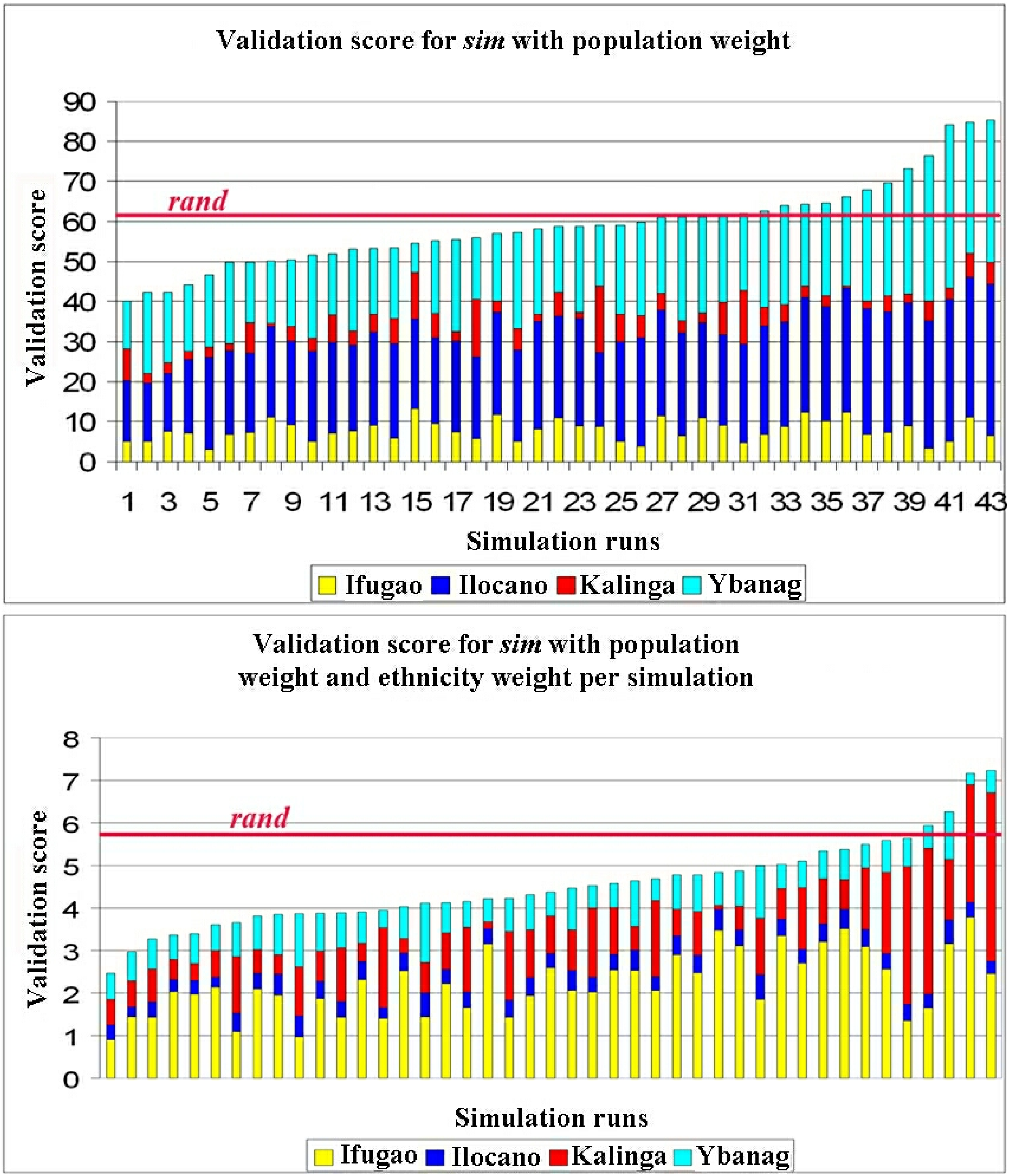
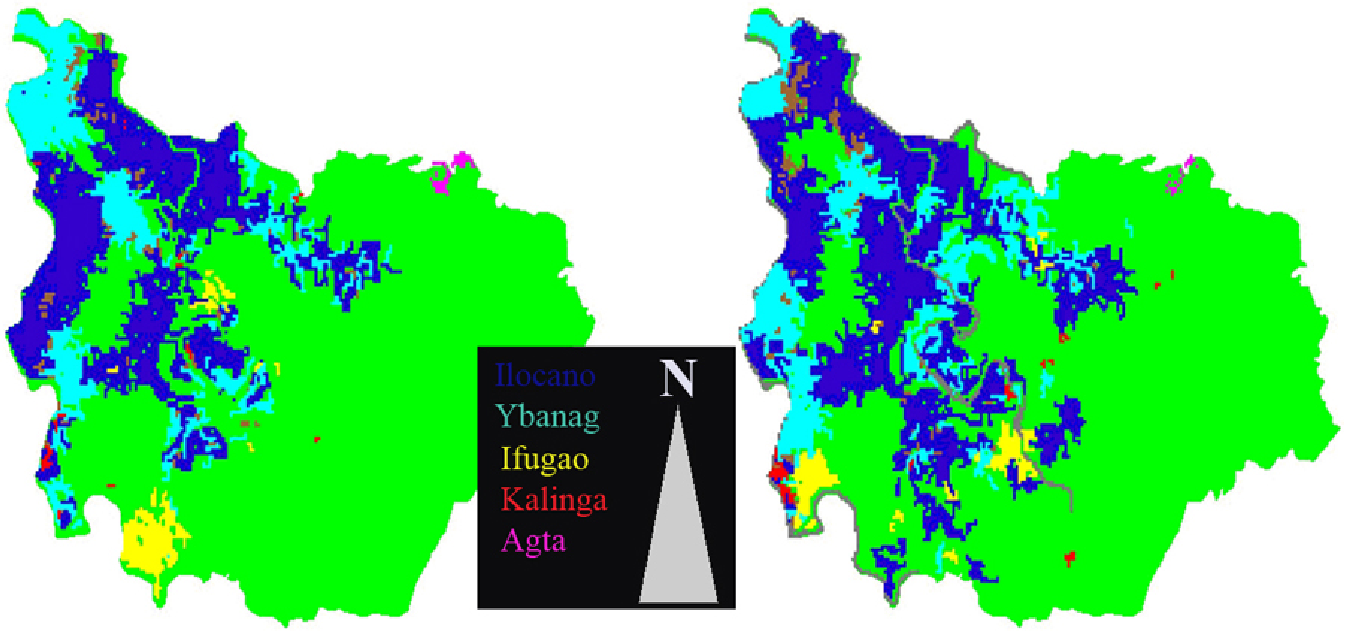


Fig. 16. Model output for simulation year 1999 with the best validation score with the population weight applied (left; batch 5, simulation ID = 1563327) and with both the population and ethnic weights applied (right; batch 5, simulation ID = 3102335).



model could possibly be improved if, during parameterization and calibration, it focused on these barangays, especially during these simulation years.

Moreover, it is quite likely that another model, in which the immigration of the various ethnicities is calibrated differently, might prove to score even better results than the current model. Future work also requires validation techniques that place more emphasis on the analysis of the emerging spatial patterns.

Stochastic initialization and path dependency

The stochastic variation in the demographic potential option nodes (PONs) GettingChildren, Marrying, and Dying (Table 7) had a large impact on the outcomes per model run. It caused large differences among the total simulated population sizes per run. In several runs, the Kalingha were almost extinct by 1999, whereas, in other runs, their population stabilized. Also, the ethnic proportions per run might differ significantly. This stochastic variation made the Kalingha population susceptible to extinction in a model run. This corresponds nicely with empirical findings; a respondent spoke about the "... difficulties of the Kalingha to plant and harvest [after World War II] because there were not enough children to help; in those days the Kalingha elders were even talking themselves about [the] extinction of their people."

The random ordering of immigrating households per year, combined with the random search direction in the settling PONs, especially during the early stages of a simulation, also had a large impact on the outcome of a model run. Once a household of a certain ethnic group had opened a new settlement, its ethnic group would flock together and other ethnicities were repulsed from that area. Hence, these stochastic elements made the model highly path-dependent and realistic. As in reality, there was no underlying factor explaining the ethnicity of the immigrating households other than word of mouth. The real ethnic immigration pattern was similar to the model: highly stochastic, causing path dependency. This is supported by a comment from

Table 6. Overall results of the MameLuke settlement model dynamic analysis. The sitio is the original settlement location.

Parameter and level		Rate of river bank occupation	Area of sitio	Degree of spatial ethnic integration
Ethnic push	High	Fast	Large	Low
	Low	Slow	Small	High
Desired land size	High	Fast	Small	Low
	Low	Slow	Large	High
Actor’s view	High	Slow	Large	Low
	Low	Fast	Small	High

one of the respondents during the interview: “How I wish that we [the Ybanag] had come some years earlier, so that we could go to the fertile places of the Ilocano.”

Stakeholder interactions

Having a starting point in a modeling exercise outside the scientific realm forces and allows the researcher to ask scientific questions that are not embedded in mainstream science. Such new questions lead to new insights and entry points for discussion. During the model construction, the generic framework proved to be very flexible and useful. New insights from primary data analysis, calibration, and stakeholder comments were easily integrated into new models. Furthermore, the way the framework was set up proved to be clear and transparent for the stakeholders. They could easily understand what and how the actors and their decisions were modeled; hence, the framework allowed for fruitful discussions, resulting in a better representation of the stakeholders’ views.

In stakeholder meetings, the dynamic model outcomes delivered new entry points in discussing a highly sensitive policy issue. The San Mariano government is aware of the increasing land scarcity because of high population growth, but never directly related the ethnic and spatial differences in population growth and land scarcity. One of the relevant outcomes, i.e., that population growth has a spatial pattern, was, in their words, an eye-opener. Their analyses were always aimed at the municipality level and never accounted for differences among ethnicities and barangays closer to or further from the town centre or forest edge. Furthermore, the model outcomes and stakeholder meetings raised renewed awareness from which the issue could be better quantified.

Another striking outcome of the modeling exercise, especially during the calibration, was the idea that a large number of families immigrated in the past. The researcher, as well as the stakeholders, estimated this immigration to be much greater than the model implied. Related to this was the misinterpretation that emigration played a significant role in population growth. Stakeholders stressed that emigration was not modeled, but played a significant part. Their feeling was that population growth was much lower than modeled. However, emigration was only 2% over 5 yr (Municipality of San Mariano 1995), and thus does not play such an important role in contemporary San Mariano barangays as was imagined.

The MameLuke settlement model outcomes indicating increasing land scarcity in combination

Table 7. Stochastic effects in the demographic potential option nodes.

Actor category	Number of people			
	Mean	Minimum	Maximum	SD
Actor	7046.70	4313	9337	894.30
Household	1536.13	1000	1978	180.75
Farm	7754.87	5021	9851	835.88
Agta	155.53	6	452	87.32
Ifugao	566.73	311	890	101.78
Ilocano	3892.19	2191	5666	698.89
Kalingha	120.64	4	505	100.84
Tagalog	424.31	233	569	69.03
Ybanag	1887.29	385	3418	665.02

with the economic reality of today suggest that the effects of population growth need to be reassessed. Because most households are farmers, whose well-being depends completely on land availability, and the opening of lands to be newly cultivated is restricted by land scarcity and policies regarding the nature park, reality demands a sound approach for future development by policy designers and implementers.

Responses to this article can be read online at:
http://www.ecologyandsociety.org/vol11/iss2/art33/responses/

Acknowledgments:

This paper is part of the project "Integrating macro-modelling and actor-oriented research in studying the dynamics of land use change in Northeast Luzon, Philippines." The project is a collaboration among Wageningen University, Nijmegen University, and Leiden University. The project is funded by the Foundation for the Advancement of Tropical Research of the Netherlands Organisation for Scientific Research (NWO-WOTRO) under project number WAI 79 70. We express our gratitude to the farmers and government officials in San Mariano municipality; the CVPED staff; Arnold Macadangdang for field assistance; Jo Ann Bagunu, Joyce Vea, Maria, Keneth, and Raquel for data encoding; two anonymous reviewers; and last but not least, Hans Bauer for helpful comments on an earlier draft.

LITERATURE CITED

Collier, N. 2001. *Repast: an extensible framework for agent simulation.* Social Science Research Computing, University of Chicago, Chicago, Illinois, USA. [online] URL: http://www.econ.iastate.edu/tesfatsi/RepastTutorial.Collier.pdf.

de Groot, W. T. 1992. *Environmental science theory: concepts and methods in a one-world, problem-oriented paradigm.* Elsevier, Amsterdam, The Netherlands.

Epstein, J. M., and R. L. Axtell. 1996. *Growing artificial societies: social science from the bottom up.* MIT Press, Cambridge, Massachusetts, USA.

Ferber, J., and O. Gutknecht. 1998. A meta-model for the analysis and design of organizations in multi-agent systems. Pages 128-135 *in* Y. Demazeau, editor. *Proceedings of the 3rd International Conference on Multi-Agent Systems (ICMAS'98)* (Paris, 1998). Institute of Electrical and Electronics Engineers, Los Alamitos, California, USA.

Fishbein, M., and I. Ajzen. 1975. *Belief, attitude, intention, and behavior: an introduction to theory and research.* Addison-Wesley, Reading, Massachusetts, USA.

Gladwin, C. H. 1989. *Ethnographic decision tree modeling.* Sage, Newbury Park, California, USA.

Griffin P. B., and A. Estioko-Griffin, editors. 1985. *The Agta of northeastern Luzon: recent studies.* San Carlos, Cebu City, Philippines.

Huigen, M. G. A. 2004. First principles of the MameLuke multi-actor modelling framework for land use change, illustrated with a Philippine case study. *Journal of Environmental Management* **72**:5-21.

Johnson, C. D., T. A. Kohler, and J. Cowan. 2005. Modeling historical ecology, thinking about contemporary systems. *American Anthropologist* **107**:96-107.

Kaimowitz, D., and A. Angelsen. 1998. *Economic models of tropical deforestation: a review.* Center for International Forestry Research, Bogor, Indonesia. [online] URL: http://www.cifor.cgiar.org/publications/pdf_files/Books/model.pdf.

Keesing, F. M. 1962. *The ethnohistory of northern Luzon.* Stanford University Press, Stanford, California, USA.

Kendall, E. A. 1999. Role modelling for agent system analysis, design, and implementation. Pages 204-218 *in First International Symposium on Agent Systems and Applications Third International Symposium on Mobile Agents (ASA/MA)* (Palm Springs, 1999). Institute of Electrical and Electronics Engineers, Los Alamitos, California, USA.

Kummer, D. M., and C. H. Sham. 1994. The causes of tropical deforestation: a quantitative analysis and case study from the Philippines. Pages 146-158 *in* K. Brown and D. W. Pearce, editors. *The causes of tropical deforestation: the economic and statistical analysis of factors giving rise to the loss of the tropical forests.* University College of London Press, London, UK.

Minar, N., R. Burkhart, C. Langton, and M. Askenazi. 1996. *The Swarm simulation system: a toolkit for building multi-agent simulations.* [online] URL: http://www.santafe.edu/projects/swarm/overview/overview.html.

Moonen, E. 1998. Land use strategies in the Sierra Madre. Pages 231-258 *in* V. T. King, editor. *Environmental challenges in South-East Asia.* Man and Nature in Asia, Volume 2. Curzon Press, Richmond, UK.

Municipality of San Mariano. 1995. *Social economic profile of San Mariano.* Municipality of San Mariano, San Mariano, The Philippines.

Municipality of San Mariano. 1997. *San Mariano census data 1997.* Municipality of San Mariano, San Mariano, The Philippines.

Municipality of San Mariano. 2002. *San Mariano census data 2002.* Municipality of San Mariano, San Mariano, The Philippines.

National Statistics Office (NSO). 1995. *The Philippine yearbook, census 1995*. Sto. Niño Catholic House, Manila, The Philippines.

National Statistics Office (NSO). 2001.*The Philippine yearbook, census 2000*. Sto. Niño Catholic House, Manila, The Philippines.

Overmars, K. P., and P. H. Verburg. 2005. Analysis of land use drivers at the watershed and household level: linking two paradigms at the Philippine forest fringe. *International Journal of Geographical Information Science* **19**:125-152.

Palo, M. 1994. Population and deforestation. Pages 42-56 *in* K. Brown and D. W. Pearce, editors. *The causes of tropical deforestation: the economic and statistical analysis of factors giving rise to the loss of the tropical forests.* University College of London Press, London, UK.

Persoon, G. A., T. Minter, B. Slee, and C. van der Hammen. 2004. *The position of indigenous peoples in the management of tropical forests.* Tropenbos

Series 23. Tropenbos International, Wageningen, The Netherlands.

Pfaff, A. S. P. 1999. What drives deforestation in the Brazilian Amazon? Evidence from satellite and socioeconomic data. *Journal of Environmental Economics and Management* **37**:26-43. [online] URL: http://www.columbia.edu/~ap196/jeem.pdf.

Rao, A. S., and M. P. Georgeff. 1995. BDI agents: from theory to practice. Pages 312-319 *in Proceedings of the First International Conference on Multi-Agent Systems (ICMAS'95)* (San Francisco, 1995). Institute of Electrical and Electronics Engineers, Los Alamitos, California, USA.

Rouchier, J., O. Barreteau, F. Bousquet, and H. Proton. 1998. Evolution and co-evolution of individuals and groups in environment. Pages 254-269 *in* Y. Demazeau, editor. *Third International Conference on Multi-Agent Systems (ICMAS'98)* (Paris, 1998). Institute of Electrical and Electronics Engineers, Los Alamitos, California, USA.

Rudel, T., and J. Roper. 1997. The paths to rain forest destruction: crossnational patterns of tropical deforestation, 1975–1990. *World Development* **25**:53-65.

Rudel, T. K. 1989. Population, development, and tropical deforestation: a cross-national study. *Rural Sociology* **54**:327-338.

Schelling, T. S. 1971. Dynamic models of segregation. *Journal of Mathematical Sociology* **1**:143-186.

Tiffen, M. 1995. Population density, economic growth and societies in transition: Boserup reconsidered in a Kenyan case-study. *Development and Change* **26**:31-66.

Tiffen, M., M. Mortimore, and F. Gichuki. 1994. *More people, less erosion: environmental recovery in Kenya.* John Wiley and Sons, Chichester, UK.

van den Top, G. M. 1998. *The social dynamics of deforestation in the Sierra Madre, Philippines.* Dissertation. Centre of Environmental Science, Leiden University, Leiden, The Netherlands.

van Weerd, M., and J. van der Ploeg. 2004. A new future for the Philippine crocodile, *Crocodylus mindorensis*. *Sylvatrop* **13**:31-50.

van Weerd, M., and R. Hutchinson. 2004. Observations of Isabela Oriole *Oriolus isabellae* in the Sierra Madre, Luzon, Philippines, with descriptions of the call. *Forktail* **20**:133-135.

Verburg, P. H., K. P. Overmars, and N. Witte. 2004. Accessibility and land-use patterns at the forest fringe in the northeastern part of the Philippines. *Geographical Journal* **170**:238-255.

Verburg, P. H., W. T. de Groot, and A. Veldkamp. 2003. Methodology for multi-scale land use change modeling: concepts and challenges. Pages 17-51 *in* A. J. Dolman, A. Verhagen, and C. A. Rovers, editors. *Global environmental change and land use.* Kluwer, Dordrecht, The Netherlands.

APPENDIX 1. Model source code. Full installation and model run instructions can be found in the readme.txt file.

Download model (File type: ZIP archive containing java and txt files; File size, unzipped: 5,864,183 bytes)

APPENDIX 2. MameLuke Settlement model run.

Download or run model (File type: MPG; File size: 5,777,412 bytes)